

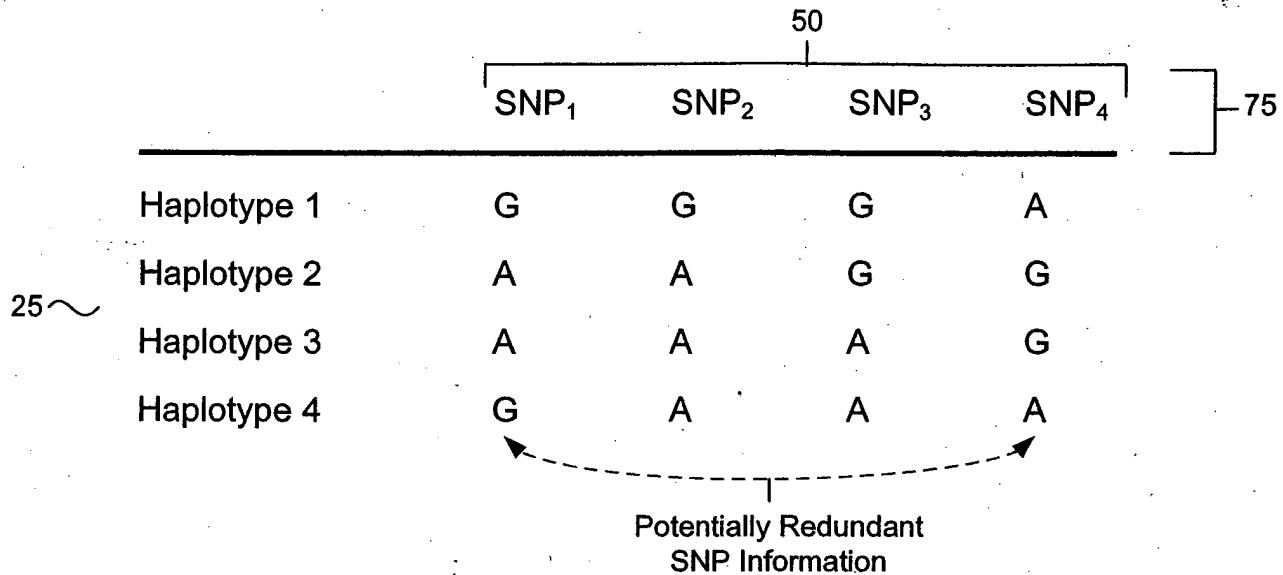
Exemplary haplotype diversity distribution (8 SNPs / 84Kb)

Potential SNP/Haplotype Combinations
 $2^{10} = 1024$

<u>Observed Haplotypes</u>	<u>Population Distribution (Frequency)</u>
1. GGACAACC	283 (83.2%)
2. AATTCGGG	40 (11.8%)
3. GATTAGCC	2 (0.6%)
4. GGTCAAGC	2 (0.6%)

FIGURE 1

Exemplary Haplotype/SNP Allele State Matrix



The diagram illustrates a matrix of haplotypes and SNPs. At the top, a horizontal bar spans four columns labeled **SNP₁**, **SNP₂**, **SNP₃**, and **SNP₄**. Above this bar is the number **50**, and to the right is a bracket spanning the four columns with the number **75**. Below the bar, four rows represent **Haplotype 1** through **Haplotype 4**, each with a corresponding column of alleles (G, G, G for Haplotype 1; A, A, G for Haplotype 2; A, A, A for Haplotype 3; G, A, A for Haplotype 4). A bracket labeled **Potentially Redundant SNP Information** spans the last two columns (**SNP₃** and **SNP₄**) for all four haplotypes. Two dashed arrows point from this bracket down to two separate matrices below.

	SNP₁	SNP₂	SNP₃
Haplotype 1	G	G	G
Haplotype 2	A	A	G
Haplotype 3	A	A	A
Haplotype 4	G	A	A

	SNP₁	SNP₂	SNP₃
Haplotype 1	G	G	G
Haplotype 2	A	A	G
Haplotype 3	A	A	A
Haplotype 4	G	A	A

	SNP₂	SNP₃	SNP₄
Haplotype 1	G	G	A
Haplotype 2	A	G	G
Haplotype 3	A	A	G
Haplotype 4	A	A	A

FIGURE 2

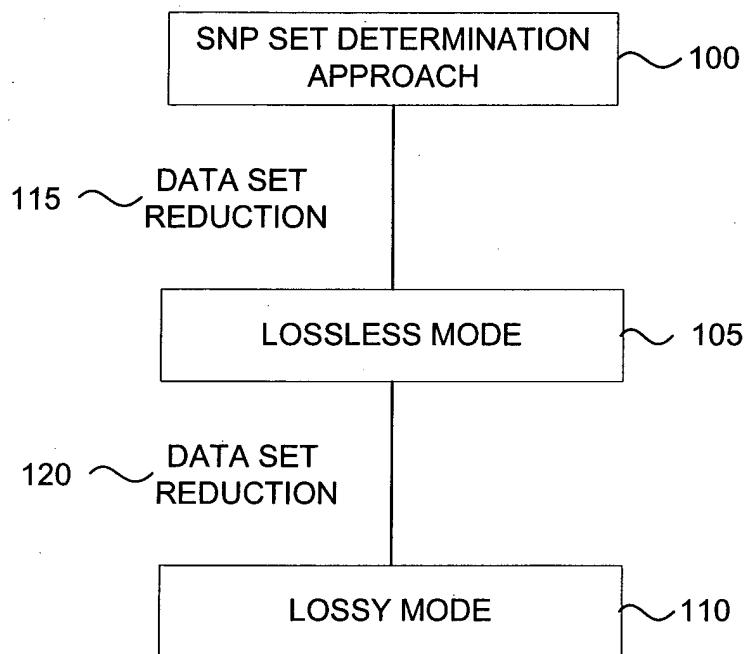
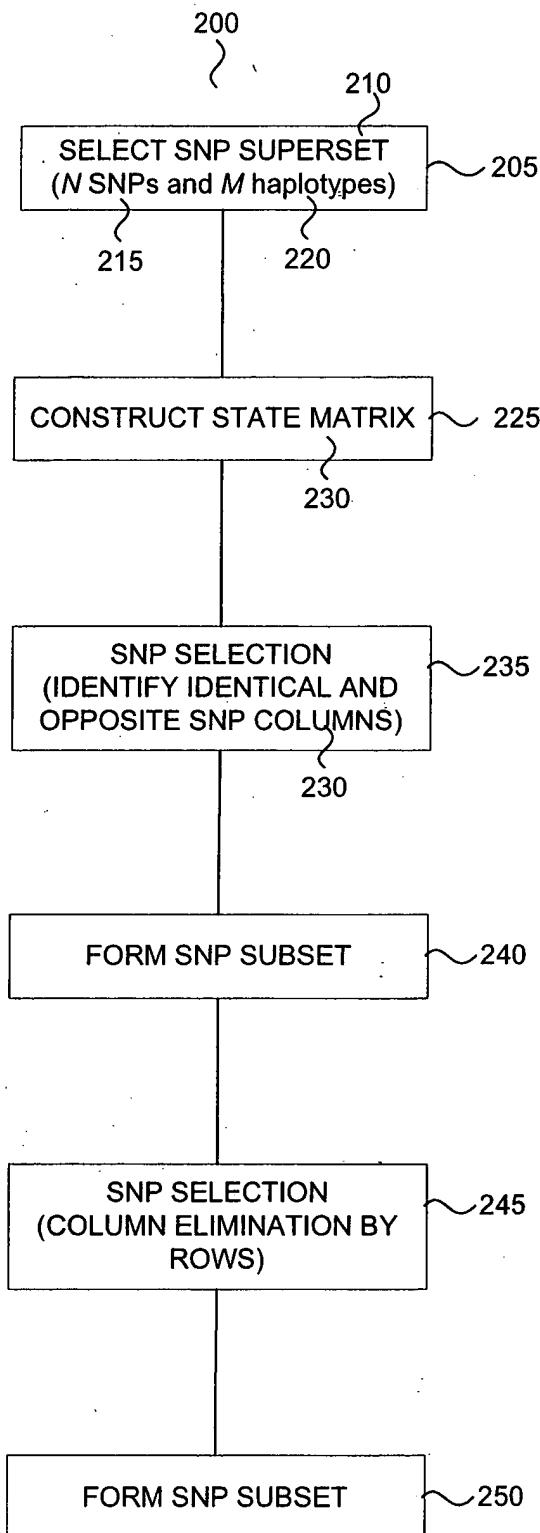


FIGURE 3



Haplotype/SNP Allele State Matrix

	SNP ₁	SNP ₂	SNP ₃	SNP ₄	SNP ₅
Haplotype 1	G	G	G	A	G
Haplotype 2	A	A	G	G	A
Haplotype 3	A	A	A	G	A
Haplotype 4	G	A	A	A	A

FIGURE 4B

FIGURE 4A

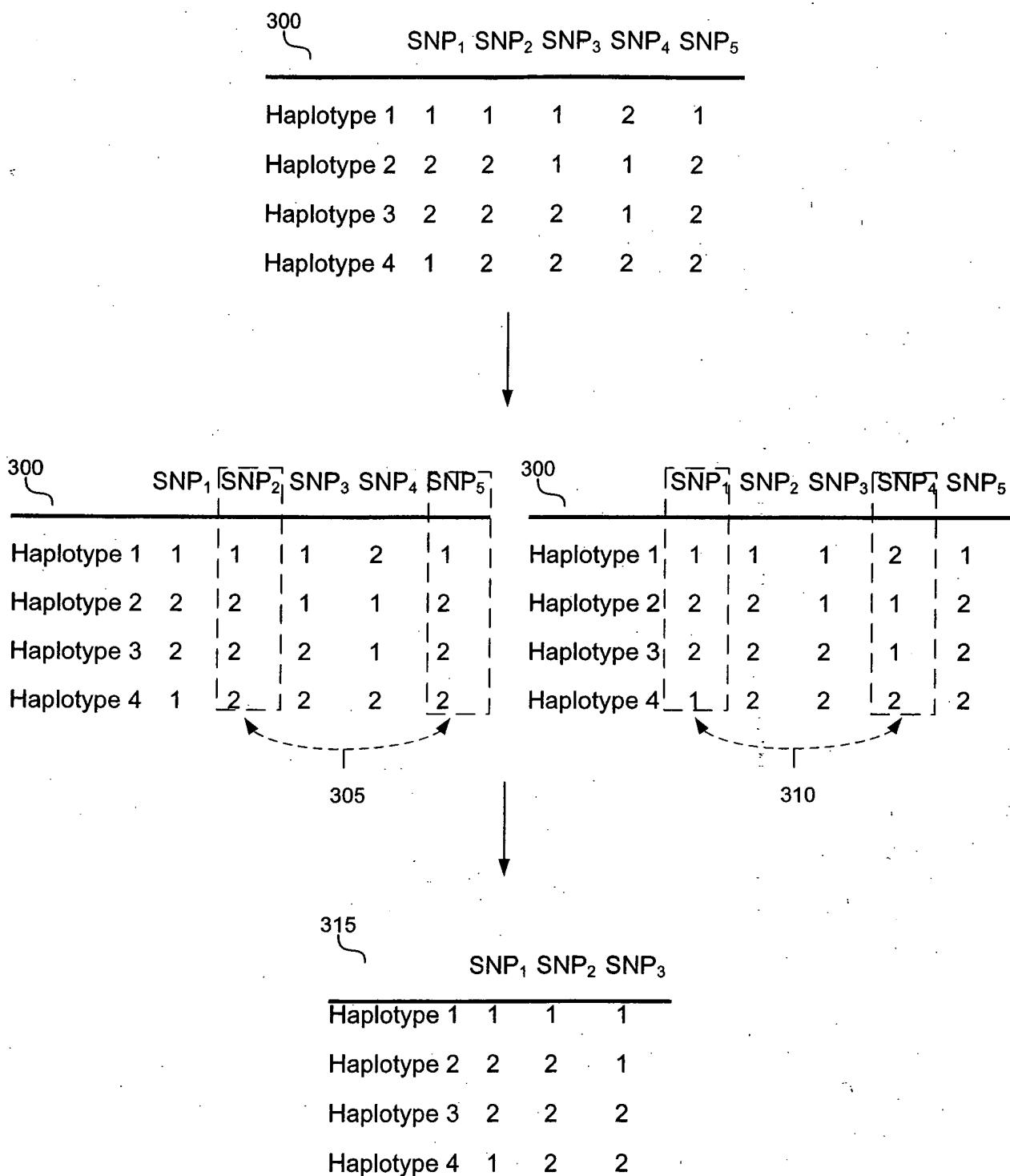


FIGURE 5A

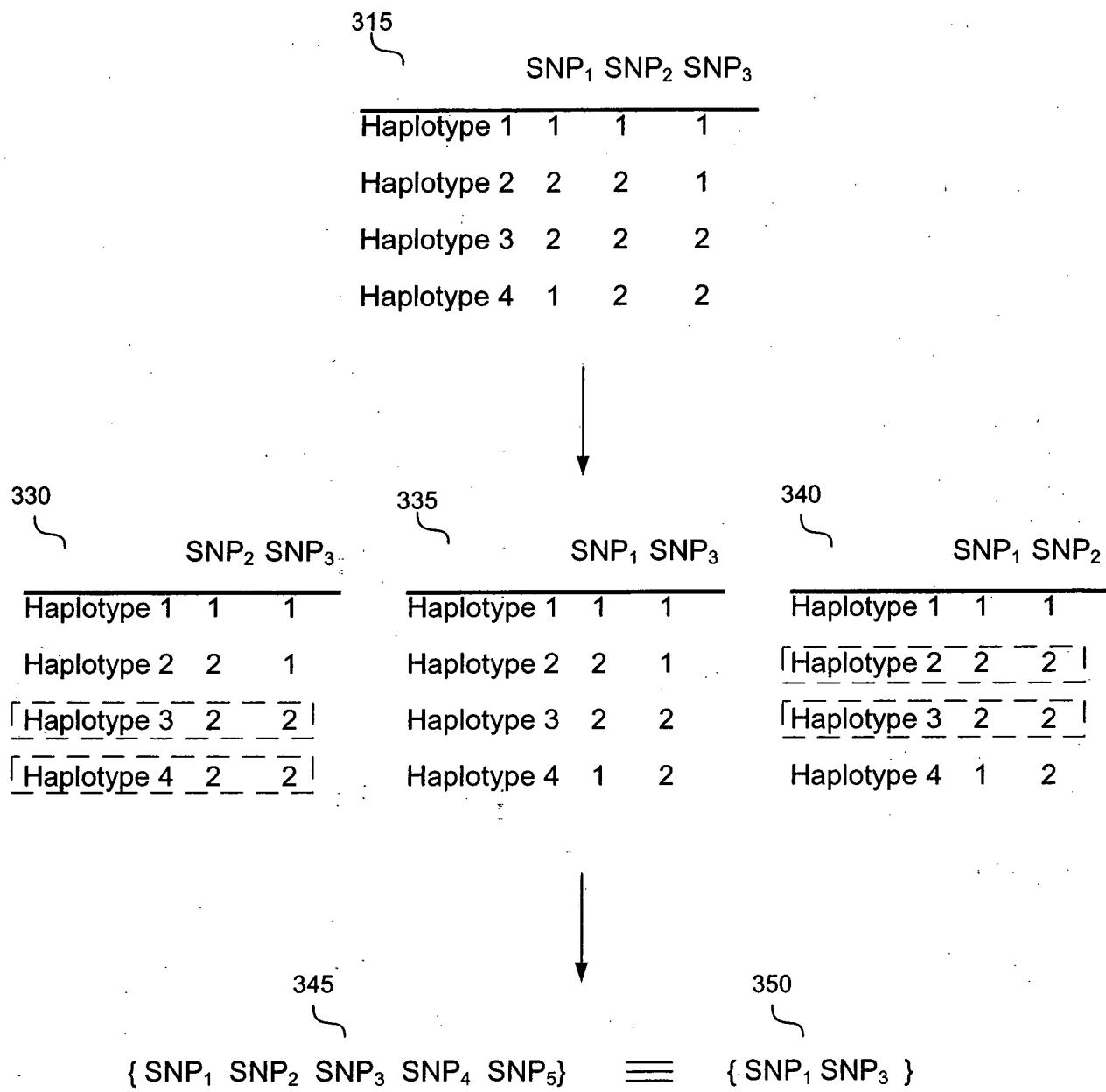


FIGURE 5B

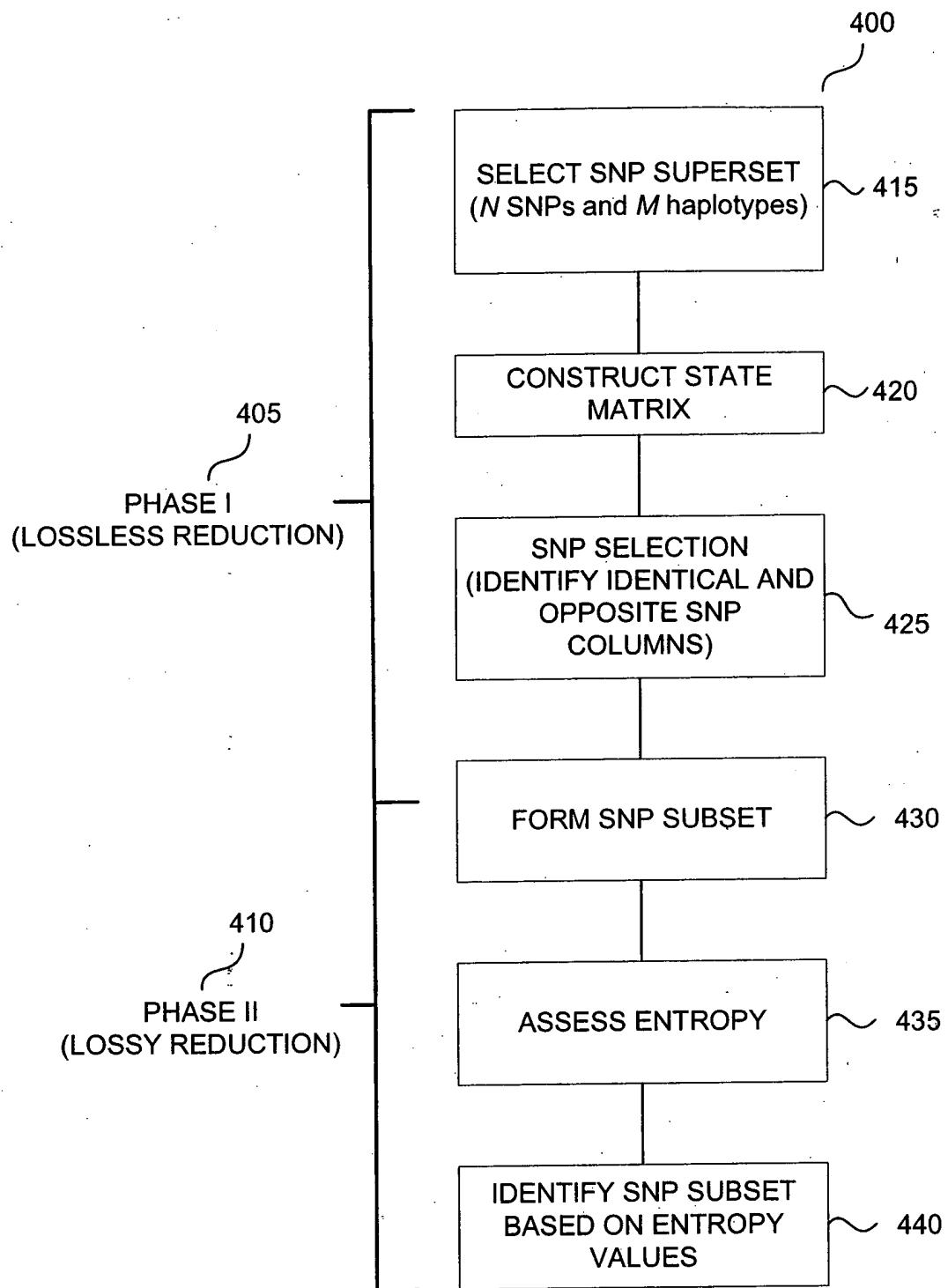


FIGURE 6A

ORIGINAL STATE MATRIX / SNP SUPERSET

455 450

M ↓ "A"

Haplotype Number	P	SNP No.																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0.1136	1	1	1	1	1	2	1	2	1	2	1	1	2	2	2	1	2
2	0.4318	1	1	1	1	1	2	1	2	1	2	1	1	2	2	2	2	2
3	0.0114	1	1	1	2	2	1	2	1	2	1	2	2	1	1	1	1	1
4	0.0454	1	2	2	1	1	2	1	2	1	2	1	1	2	2	2	1	1
5	0.0454	2	1	1	1	1	2	1	2	1	2	1	1	2	2	2	1	2
6	0.0118	2	2	2	2	2	1	2	1	2	1	2	1	1	1	1	1	1
7	0.3292	2	2	2	2	2	1	2	1	2	1	2	2	1	1	1	1	1
8	0.0114	2	2	2	2	2	1	2	1	2	2	1	1	1	1	1	1	1

454 N → 460

FIGURE 6B

RESULTING STATE MATRIX FOLLOWING
 APPLICATION OF LOSSLESS APPROACH

455 465

Haplotype Number	P	SNP ₁	SNP ₂	SNP ₄	SNP ₁₀	SNP ₁₂	SNP ₁₆	SNP ₁₇
1	0.1136	1	1	1	2	1	1	2
2	0.4318	1	1	1	2	1	2	2
3	0.0114	1	1	2	1	2	1	1
4	0.0454	1	2	1	2	1	1	1
5	0.0454	2	1	1	2	1	1	1
6	0.0118	2	2	2	1	1	1	1
7	0.3292	2	2	2	1	2	1	1
8	0.0114	2	3	2	2	2	1	1

FIGURE 6C

IDENTIFICATION OF HAPLOTYPE DIVERSITY
De La Vega et al.
Appl. No.: Unknown *Atty Docket: ABIOS.048A*

No. of SNPs (<i>k</i>)	No. of Combinations ***	Optimal Set of <i>k</i> SNPs	Haplotype Distribution Resulting from the Optimal SNP Set	Resulting Entropy (<i>H</i>) (bits)
7	8	{SNP ₁ , SNP ₂ , SNP ₄ , SNP ₁₀ , SNP ₁₂ , SNP ₁₅ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
6	28	{SNP ₁ , SNP ₂ , SNP ₁₀ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
5	56	{SNP ₁ , SNP ₁₀ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
4	70	{SNP ₁ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.341)	1.9631
3	56	{SNP ₁ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.057, 0.045, 0.352)	1.8475
2	28	{SNP ₁₂ , SNP ₁₆ }	(0.216, 0.432, 0.397)	1.5311
1	8	{SNP ₁₆ }	(0.5682, 0.4318)	0.9865

FIGURE 6D

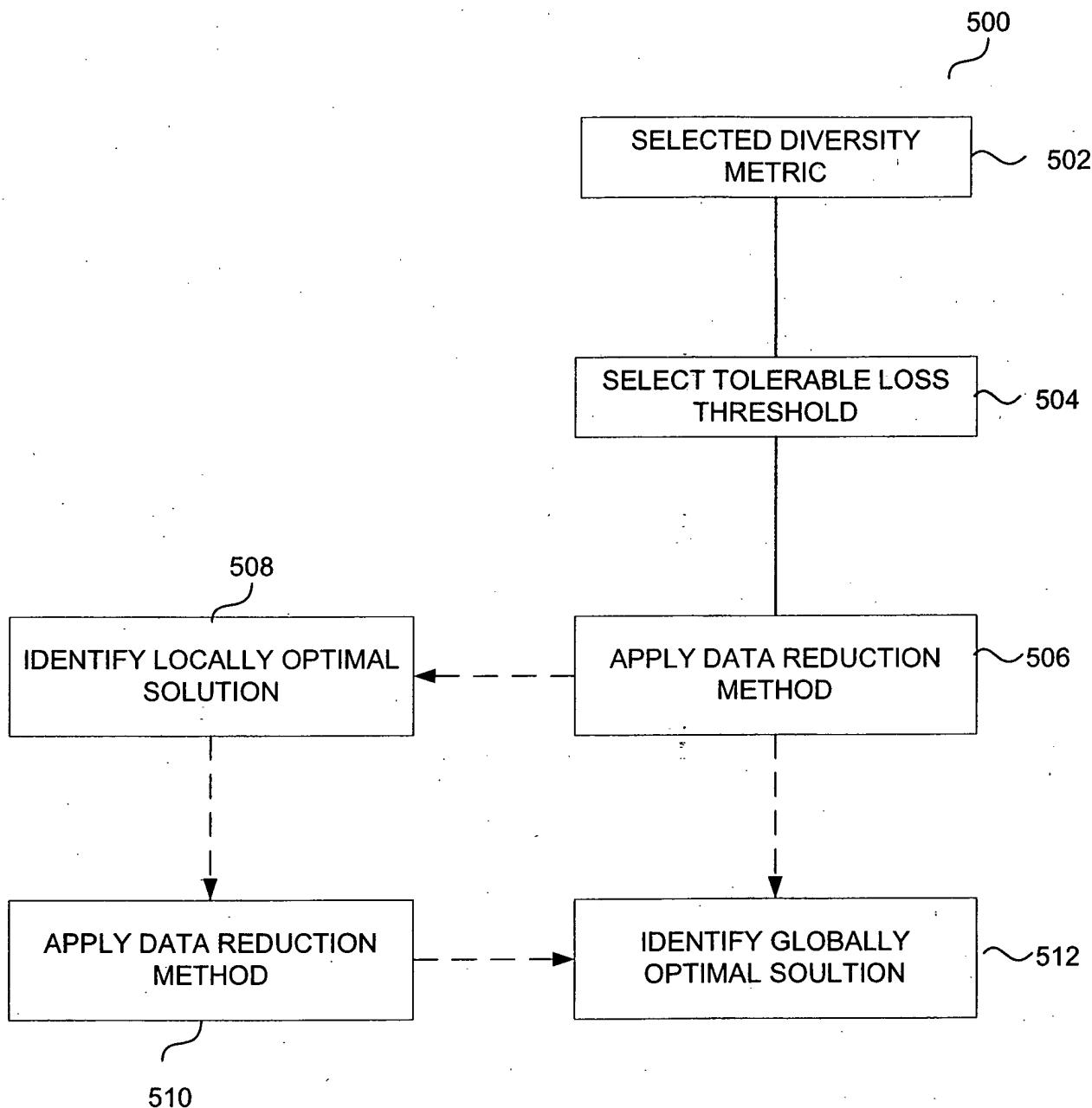


FIGURE 7

				557		557	
				P			
		SNP ₁	SNP ₂	SNP ₃	SNP ₄		
Hap1		1	1	1	2	0.63	
Hap2		2	2	1	1	0.22	
Hap3		2	2	2	1	0.07	
Hap4		1	2	2	2	0.08	
						558	
						Remaining Info (H)	
No. of SNPs		No. of Combinations $4! / (K! (4-K)!)$		Optimal Set of SNPs			
(K)				{SNP ₁ SNP ₂ SNP ₃ SNP ₄ }		1.461	
4		1		{SNP ₁ SNP ₂ SNP ₃ }		1.461	
3		4		{SNP ₁ SNP ₃ }		1.461	
559 ~ 2		6		{SNP ₂ }		0.951	
1		4					

FIGURE 8

Chr.	Pop.	Total No. of SNPs	Mean Spacing Between SNPs (bp)	546	548	550	552	556	Mean Min. SNP per Block	
									No. of Haplotype Blocks	Mean Block Size (bp)
6	A	2,504	24,386	10,840	646	23,000	3,88	2,94	2,44	2,33
	C	4,009	23,694	10,630	883	34,000	4,54	2,86	2,47	555
21	A	955	12,424	7,382	242	14,933	3,95	2,92	2,39	540
	C	1,555	11,921	7,031	336	21,032	4,63	2,88	2,32	545
22	A	1,405	10,041	6,035	350	13,714	4,01	2,99	2,47	535
	C	1,783	9,080	7,760	417	17,505	4,28	2,81	2,27	

FIGURE 9

IDENTIFICATION OF HAPLOTYPE DIVERSITY
De La Vega et al.
Appl. No.: Unknown Atty Docket: ABIOS.048A

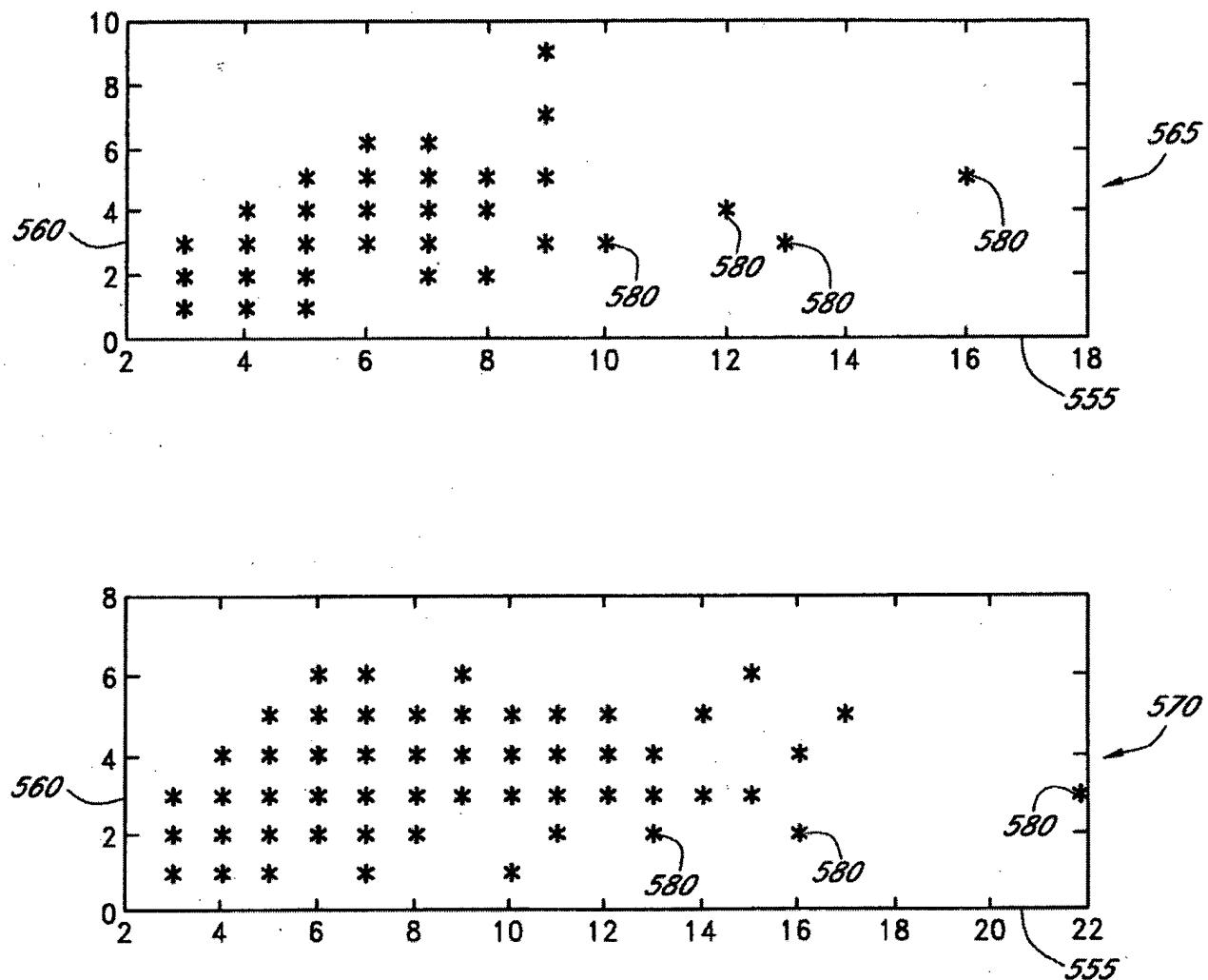


FIGURE 10